

SVEUČILIŠTE U ZAGREBU
FAKULTET ELEKTROTEHNIKE I RAČUNARSTVA

Diplomski seminar 2 – izvještaj

**Utjecaj augmentacije podataka na točnost
klasifikacije koristeći konvolucijsku neuronsku
mrežu**

Antun Herkov

Zagreb, svibanj 2022.

Sadržaj

Uvod.....	3
1. Generalna problematika.....	4
2. Baza podataka.....	5
1. Prilagodba baze podataka.....	5
3. Metodologija.....	6
1. Arhitektura mreže.....	6
2. Učenje modela.....	6
1. Transformacije.....	6
2. Funkcija pogreške.....	7
3. Metrika točnosti.....	7
4. Eksperimenti i rezultati.....	7
1. Tri klase, bez augmentacije.....	7
2. Tri klase, s augmentacijom.....	8
3. Pet klasa, bez augmentacije.....	8
4. Pet klasa, s augmentacijom.....	9
Zaključak.....	9
Literatura.....	10

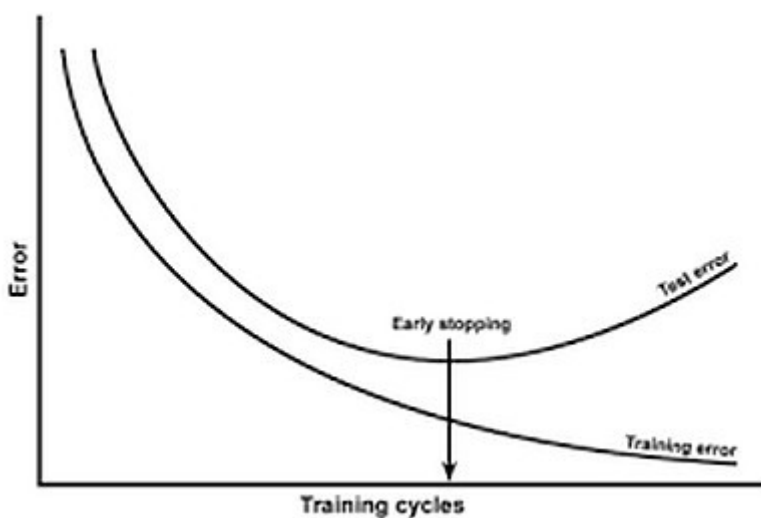
Uvod

Poznato je da veća baza podataka rezultira boljim modelima dubokoga učenja no stvaranje velikih baza podataka se nerijetko pokazuje izazovom zbog potrebe za ručnim prikupljanjem i označavanjem podataka.

Ovo se pokazuje posebno izazovnim u stvaranju većih skupova podataka u području medicine i medicinskoga snimanja. Takvi izazovi su motivacija iza brojnih istraživanja na temu augmentacije podataka.

Konvolucijske neuronske mreže imale su vrlo dobre rezultate u području računalnoga vida. Nažalost, takve mreže se oslanjaju na velike skupove podataka kako bi se izbjegla pojava prenaučenosti.

Prenaučenost je pojava u raznim područjima umjetne inteligencije gdje se model nauči na podatke iz skupa podataka za učenje te gubi svojstvo generalizacije – mogućnost zaključivanja na temelju neviđenih podataka.



Slika 1: područja podnaučenosti i prenaučenosti modela

1. Generalna problematika

Ipak, postoje područja gdje nemamo velike skupove podataka, bilo to iz razloga da podaci ne postoje ili da je označavanje istih skup postupak. Nekolicina radova se fokusira na problem augmentacije podataka – potencijalno rješenje problema limitiranoga skupa podataka.

Augmentacija podataka je skup različitih tehnika čiji je cilj povećati skup dostupnih podataka za učenje i eventualno kvalitetu postojećih podataka tako da se iz njih mogu naučiti uspješniji modeli dubokoga učenja. Neke od općenito korištenih tehnika uključuju geometrijske transformacije, promjene u prostoru boja, *kernel* filtere, korištenje GAN-ova (engl. *generative adversarial networks*) itd.

U usporedbi s ostalim metodama borbe protiv prenaučivosti u strojnom učenju, augmentacija podataka pristupa tome problemu u samome skupu za učenje.

Početna pretpostavka je da se augmentacijom može iz originalnog skupa podataka izlučiti više informacija nego što je prethodno dostupno. Takva vrsta augmentacija povećava skup podataka ili izmjenama ili preuzorkovanjem.

Prvi tip transformira podatke tako da je njihova oznaka ostala ista (geometrijske transformacije i promjene boja, nasumična brisanja i sl.) dok drugi tip stvara sintetičke podatke i dodaje ih skupu podataka za učenje (GAN-ovi). Ovakvi načini augmentacije nisu međusobno isključivi - moguće je npr. primijeniti geometrijske transformacije nad kreiranim sintetičkim podacima kako bi dodatno augmentirali naš skup podataka.

2. Baza podataka

1. Prilagodba baze podataka

Početna baza podataka sastojala se od 1,000 fotografija polica s mnoštvom kutija čajeva. Za prilagodbu baze podataka problemu kojime sam se bavio te korištenoj arhitekturi mreže bilo je potrebno odabrati proizvoljan broj klasa odnosno vrsti čajeva koje će mreža prepoznavati.

U slučaju provedenih eksperimenata to su bile klase "AppleCinnamonTeaBox", "BlueberryTeaBox", "GreenTeaBox", "AroniaTeaBox" i "CamomileTeaBox".

Koristeći prethodno detektirane labele, bilo je potrebno izrezati RGB slike samo odabranih klasa. Iz dataseta su također izbačeni svi primjeri kojima je vidljivost (polje 'visibility') bila manja od praga postavljenoga na 40.00. To polje predstavlja koliki postotak objekta je zapravo vidljiv, s obzirom da su neki objekti djelomično skriveni ostalima kutijama ili samim policama.

Podaci su zatim podijeljeni na skup za učenje, validaciju i testiranje u omjeru 70:20:10.



Slika 2: Primjer originalne slike čajeva na polici iz dataseta



Slika 3: Primjeri cropanih slika

3. Metodologija

1. Arhitektura mreže

U sklopu praktičnog dijela ovoga seminara u PyTorchu je implementirana slična arhitektura mreže koja se koristila i u radu „*The Effectiveness of Data Augmentation in Image Classification using Deep Learning*” (J. Wang, L. Perez).

Pregled slojeva arhitekture:

1. Conv2d, 16 kanala, 3×3 filteri. ReLU
2. Batch normalizacija
3. Max pooling, 2×2 filteri, 2×2 korak
4. Conv2d, 32 kanala, 3×3 filteri. ReLU
5. Conv2d, 32 kanala, 3×3 filteri. ReLU
6. Batch normalizacija
7. Max pooling, 2×2 filteri, 2×2 korak
8. Fully connected sloj s izlazom veličine 1024. Dropout
9. Fully connected sloj s izlazom veličine 3 ili 5 - ovisno o eksperimentu

2. Učenje modela

Pri učenju modela se nakon svake epohe radila validacija modela na validacijskom skupu podataka te se koristila tehnika ranog zaustavljanja učenja modela (engl. *early stopping*) u slučaju da se točnost klasifikacije na testu za validaciju zadnjih tri epohe nije poboljšala. Tijekom učenja korišten je optimizator Adam.

1. Transformacije

U data loaderu se na svakom batchu slika primjene nasumične transformacije koristeći PyTorch-ovu RandAugment opciju.



Slika 4: Primjer nasumičnih transformacija opcijom RandAugment

2. Funkcija pogreške

Modeli su ućeni optimizacijom višeklasne logistićke pogreške. Unakrsna entropija (engl. *cross entropy*) daje nam vjerojatnosnu sličnost između prave oznake (engl. *ground truth*) i predviđene vrijednosti pri trenutnom stanju mreže.

Prosjećna unakrsna entropija svih klasa definira logistićku pogrešku koja kašnjava devijaciju predviđene vjerojatnosti pripadanja pojedinoj klasi od stvarne pripadnosti nekoj od klasa.

3. Metrika toćnosti

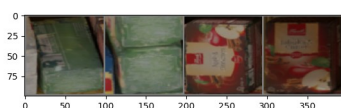
Kao metrika toćnosti odabrana je metrika toćnosti klasifikacije pojedinih slika. Konceptualno je vrlo jednostavna i izražava se u postotku slika koje su ispravno klasificirane te se pokazala kao dobar pokazatelj toćnosti kad je primijenjena na naš problem.

4. Eksperimenti i rezultati

Nad datasetom provedeno je nekoliko eksperimenata - svaki ućen do konvergencije te su u ostatku izvještaja prikazani postignuti rezultati.

1. Tri klase, bez augmentacije

Primjer primjene modela na nasumićan podskup slika iz dataseta:



GroundTruth: GreenTeaBox GreenTeaBox
AppleCinnamonTeaBox AppleCinnamonTeaBox

Predicted: GreenTeaBox GreenTeaBox
AppleCinnamonTeaBox AppleCinnamonTeaBox

Tablica 1: Tri klase, bez augmentacije

Točnost na skupu za testiranje	98%
AppleCinnamonTeaBox	99.8%
BlueberryTeaBox	98.3%
GreenTeaBox	98.3%

2. Tri klase, s augmentacijom

Primjer primjene modela na nasumičan podskup slika iz dataseta:



GroundTruth: AppleCinnamonTeaBox
 AppleCinnamonTeaBox BlueberryTeaBox
 BlueberryTeaBox

Predicted: AppleCinnamonTeaBox
 AppleCinnamonTeaBox BlueberryTeaBox
 BlueberryTeaBox

Tablica 2: Tri klase, s augmentacijom

Točnost na skupu za testiranje	99%
AppleCinnamonTeaBox	98.4%
BlueberryTeaBox	99.7%
GreenTeaBox	99.7%

3. Pet klasa, bez augmentacije

Primjer primjene modela na nasumičan podskup slika iz dataseta:



GroundTruth: BlueberryTeaBox AroniaTeaBox
 AppleCinnamonTeaBox AroniaTeaBox

Predicted: BlueberryTeaBox AroniaTeaBox
 AppleCinnamonTeaBox AroniaTeaBox

Tablica 3: Pet klasa, bez augmentacije

Točnost na skupu za testiranje	97%
AppleCinnamonTeaBox	94.0%
AroniaTeaBox	98.7%
BlueberryTeaBox	96.6%
CamomileTeaBox	98.5%
GreenTeaBox	97.6%

4. Pet klasa, s augmentacijom

Primjer primjene modela na nasumičan podskup slika iz dataseta:



GroundTruth: CamomileTeaBox GreenTeaBox
GreenTeaBox CamomileTeaBox

Predicted: CamomileTeaBox GreenTeaBox
GreenTeaBox CamomileTeaBox

Tablica 4: Pet klasa, bez augmentacije

Točnost na skupu za testiranje	98%
AppleCinnamonTeaBox	98.9%
AroniaTeaBox	97.0%
BlueberryTeaBox	99.9%
CamomileTeaBox	96.2%
GreenTeaBox	99.0%

Zaključak

Ovim seminarom nastojalo se prikazati rezultate primjene augmentacije na podatke korištene u klasifikacijskom problemu. Opisana je metodologija te dan pregled korištene arhitekture mreže. Razvijena je implementacija iste korištenjem biblioteke PyTorch pomoću koje je definirana konvolucijska neuronska mreža te maksimalno iskorištena efikasnost učenja koristeći grafičke procesore u svrhu obrade podataka.

Klasifikacija korištenjem ovakve arhitekture dala je vrlo dobre rezultate i u slučaju ulaznih podataka bez korištenja postupaka augmentacije i u slučaju s korištenjem, međutim, rezultati nad skupom podataka za testiranje su se ipak pokazali uspješniji primjenom modela koji su koristili nasumične transformacije i na taj način augmentirali skup podataka.

Literatura

- [1] Wang, J; Perez, L: The Effectiveness of Data Augmentation in Image Classification using Deep Learning, Stanford University, 2017
- [2] J. Torres. Learning Process of a Deep Neural Network. Medium, Towards Data Science (2020). Poveznica: <https://towardsdatascience.com/learning-process-of-a-deep-neural-network-5a9768d7a651>
- [3] PyTorch - open source machine learning framework. Poveznica: <https://pytorch.org/>
- [4] Comet - machine learning platform. Poveznica: <https://www.comet.ml/site/>